

Guidelines for Employers to Use a Random Sample of Employees to Estimate Average Vehicle Ridership (AVR)

**(Option Provided to Employers Exempted from Rule 2202 and using the
Employee Commute Reduction Program Option)**

**Transportation Programs
South Coast Air Quality Management District
21865 Copley Drive
Diamond Bar, CA 91765**

AUTHOR: WALDO AQUERES-LOPEZ, AIR QUALITY SPECIALIST
JANUARY 30, 1998

Table of Contents

ACKNOWLEDGMENTS	1
A. MAJOR STEPS IN USING THE SAMPLING METHOD	2
B. INTRODUCTION AND OBJECTIVES	2
C. APPLICABILITY AND GENERAL REQUIREMENTS: EXCERPTS FROM THE EMPLOYEE COMMUTE REDUCTION PROGRAM GUIDELINES	4
D. SAMPLING CONCEPTS AND TERMINOLOGY	4
1. <i>Population</i>	4
2. <i>Sampling Unit</i>	5
3. <i>Probability Sample</i>	5
4. <i>Sampling Frame</i>	5
5. <i>Representative Sample</i>	5
6. <i>Sampling Error</i>	6
7. <i>Reliability</i>	6
8. <i>Sampling Techniques</i>	6
a. Simple Random Sampling.....	6
b. Stratified Random Sampling	7
c. Systematic Sampling	8
E. REQUIRED SAMPLE SIZE	8
1. <i>Sampling Error and Reliability Criteria</i>	8
2. <i>Calculation of Sample Size</i>	9
3. <i>Non-responses</i>	10
F. THE AQMD’S RECOMMENDED SAMPLING METHOD	11
1. <i>Drawing the Sample</i>	11
2. <i>Working with Inexact Skip Intervals</i>	12
3. <i>Multi-Site Employers</i>	13
G. EXCEPTIONS TO THE AQMD’S RECOMMENDED SAMPLING METHOD	14
H. REQUESTING APPROVAL TO USE A RANDOM SAMPLE.....	14
1. <i>AQMD’s Method</i>	14
2. <i>Other than the AQMD’s Method</i>	15
I. REQUIRED SAMPLE DOCUMENTATION	15
APPENDIX I - STATISTICAL FORMULAS USED TO DETERMINE SAMPLE SIZES	18
APPENDIX II - REFERENCES	19

Acknowledgments

These sampling guidelines were developed with the support of Transportation Programs staff and many other individuals. The following advisory panel members contributed significantly to the final product by raising key issues and making practical recommendations, most of which were incorporated into the document:

Anthony Campagna, West Hollywood Area TMO

Ross Hopkins, Ross Hopkins, and Associates

Mick O'Donnell, Riverside Community College

Sammy Rake, City of Irvine

Yvonne Theriault, McDonnell Douglas Corporation

Richard Waide, Ekistics Mobility Consultants

Roger Windeler, Hughes Aircraft

Special thanks go to Philip G. Enns, Statistician of the U.S. Environmental Protection Agency, who took the time to review the report and examine the statistical formulas used in the calculation of sample sizes.

A. Major Steps in Using the Sampling Method

Employers who choose the sampling method to calculate the worksite's AVR should do the following:

1. Ascertain that the number of employees in the peak window is at least 400;
2. Determine the exact number of employees arriving to work during the designated peak window;
3. Use Table 2 to determine the appropriate sample size;
4. Select the sample of employees according to:
 - a. The AQMD's recommended method, as described in Section F; or
 - b. The employer's preferred method of selection, as explained in Section G;
5. Request approval from the AQMD to use the sampling method to calculate AVR (Section H);
6. Conduct survey; and
7. Submit sample documentation with your triennial plan or annual analysis, as specified in Section I.

Employers who are planning to use the AQMD's recommended sampling method may skip the more technical sections of the guidelines and refer to Sections E and F which discuss the required sample sizes and the AQMD's recommended sampling design.

B. Introduction and Objectives

According to the *Employee Commute Reduction Program Guidelines* adopted on December 8, 1995, employers with a minimum of 400 employees reporting at the worksite between 6:00 and 10:00 a.m. can use a random sample to estimate average vehicle ridership (AVR). Approximately, 11% (611) of all worksites subject to Rule 2202 have at least 400 employees reporting to work between 6:00 and 10:00 a.m. (Table 1). There are a few worksites with as many as 22,000 employees reporting to work at this time of the day. Thus, sampling will definitely reduce AVR survey costs for large employers which have decided to use the employee commute reduction program as an alternative to Rule 2202.

This document sets forth guidelines containing general and specific requirements to help employers in developing an adequate sampling procedure to estimate AVR. Topics addressed include major steps required in using the sample method; general criteria for employers to follow in using a sampling method; concepts and terms used in sampling; required sample size for employers with different number of employees; the AQMD's recommended, or preferred, sampling method; exceptions to the AQMD's recommended approach; documentation required to request AQMD's approval of employer's proposed

sampling method; and sampling documentation required to be submitted to evaluate the validity of the sample.

The statistical concepts, terms, and sampling designs addressed in these guidelines provide background material for understanding and justifying the AQMD's recommended sampling design or choice of sampling method. Throughout this report, emphasis is placed on the mechanics of sampling selection, with little discussion of sampling theories and statistical calculations. Several works on sampling have been used in preparing these guidelines and they are listed in Appendix II. The works of Babbie (1973), Kish (1965), and Yamane (1967) in particular, have been extremely useful. Statistical formulas are included in Appendix I.

Table 1
Number and percent of worksites and employees subject to Rule 2202 by
number of employees in the 6 a.m. to 10 a.m. window

Window Employees		Worksites	Percent of Sites	Cumulative Percent *	6-10 Employees	Percent Employees**	Cumulative Percent
From	To						
0	99	1,569	29.0%	100.0%	78,450	5.6%	5.6%
100	199	2,182	40.4%	71.0%	327,300	23.4%	29.0%
200	299	717	13.3%	30.6%	179,250	12.8%	41.8%
300	399	323	6.0%	17.3%	113,050	8.1%	49.9%
400	499	180	3.3%	11.3%	81,000	5.8%	55.7%
500	599	95	1.8%	8.0%	52,250	3.7%	59.5%
600	699	68	1.3%	6.2%	44,200	3.2%	62.6%
700	799	47	0.9%	5.0%	35,250	2.5%	65.1%
800	899	32	0.6%	4.1%	27,200	1.9%	67.1%
900	999	27	0.5%	3.5%	25,650	1.8%	68.9%
1,000	1,999	98	1.8%	3.0%	147,000	10.5%	79.4%
2,000	2,999	28	0.5%	1.2%	70,000	5.0%	84.4%
3,000	3,999	14	0.3%	0.7%	49,000	3.5%	87.9%
4,000	4,999	2	0.0%	0.4%	9,000	0.6%	88.6%
5,000	5,999	3	0.1%	0.4%	16,500	1.2%	89.8%
6,000	6,999	4	0.1%	0.3%	26,000	1.9%	91.6%
7,000	7,999	3	0.1%	0.2%	22,500	1.6%	93.2%
8,000	8,999	3	0.1%	0.2%	25,500	1.8%	95.1%
9,000	9,999	2	0.0%	0.1%	19,000	1.4%	96.4%
10,000	or more	5	0.1%	0.1%	50,000	3.6%	100.0%
Total		5,402	100.0%		1,398,100	100.0%	

(1) Percentages accumulated from bottom to top.

March 1995

(2) Percentages accumulated from top to bottom.

C. Applicability and General Requirements: Excerpts from the Employee Commute Reduction Program Guidelines

Employers with a minimum of 400 employees reporting at the worksite during the window used for calculating AVR have the option of determining AVR by a random sample method. The random sample method must comply with all of the following criteria:

- Members of the sample must be selected in a probability basis (i.e., random selection) that assures that each population member is given an equal chance of selection;
- All employees reporting in the window for calculating AVR must be considered as the relevant population from which the sample is drawn;
- The sample must measure all potential commute modes for employees arriving at the worksite during the window and shall account for all employees not arriving at the worksite during the window due to compressed day off, vacation, sick leave, and other (e.g., maternity leave, bereavement leave, etc.);
- Any employee designated for the random sample who does not respond to the survey must be counted as a solo driver; a minimum of 60% survey response rate must be attained;
- The sample size must be determined with the AQMD's approval of sampling method;
- The data shall be stored and available for inspection for three years;
- Any data submitted in a computer readable disk must be compatible with the AQMD's software and must be able to be entered into the AQMD's system;
- The random sample survey must be taken not more than 6 months prior to submittal of the triennial program or annual analysis; and
- The random sample method must receive written approval from the AQMD prior to administration of the survey;

D. Sampling Concepts and Terminology

1. Population

Population or universe is defined as the totality of elements, things, or people under consideration. For purpose of the employee random sample, the population consists of all employees reporting to work within the established window. For example, if the established window is from 6:00 a.m. to 10:00 a.m., the population would include all employees reporting to work during that time. Depending upon employment fluctuations, this number may vary from year to year.

2. Sampling Unit

A sampling unit is the element, item, person, or thing considered for selection in the sample. Thus, the population is made up of sampling units. For an employer electing to implement an employee commute reduction program, the sampling units are the employees arriving to work during the regulatory window. For example, an employer who has 1,000 employees arriving to work during the established time window will have a population of 1,000 employees or 1,000 sampling units.

3. Probability Sample

A probability sample consists of selecting a portion (or sample) of the population so that each unit or member of the population has an equal chance of being included in the sample. A random selection process is essential to draw a probability sample. Thus, a random sample is a probability sample by definition. A random sample also allows the employer to make generalizations from a small sample to a wider population of employees.

In selecting a random sample, a table of random numbers or a computer program must be used to select the sample. This approach prevents bias on the part of the person choosing the sample. The random selection of sampling units makes possible the calculation of sampling error and the quantification of the reliability or degree of accuracy of the sample (Babbie, 1973).

4. Sampling Frame

The sampling frame is the actual list of sampling units or a listing of employee names from which the employer can select the sample. To pick a random sample of employees, the first thing that an employer has to do is to compile a list of employees reporting to work during the established window. This list is usually available through personnel records and can be arranged in alphabetical order by employee name.

5. Representative Sample

A representative sample ensures that the most relevant characteristics of the population (e.g., mode split, type of transportation used to work, occupation, age group, etc.) are represented in the sample in roughly the same proportion they are found in the population (Kerlinger, 1973). While a random or probability sample may not be representative of all the characteristics of the population, it must be fairly representative of those characteristics that may be relevant to what is being measured or estimated. In choosing a sample, employers must ensure that the sample is random and representative of the driving habits of window employees.

The sampling frame may include other employee characteristics which may help pick a representative sample. These characteristics may include employee occupation or job category. Employee occupation is useful because of its association with ridesharing.

Managers and professional employees, for example, appear to have lower carpool propensities than blue collar workers (Stevens, 1990). Thus, in selecting the sample, it is critical that the various employee occupations be well represented. Other characteristic may include education, age, or any other specific employee attribute which may help yield a representative sample of employees' modal share or driving habits.

6. Sampling Error

Sampling error is the difference between the sample value and the population value. Margin of error is a more common term for sampling error. The margin of error or precision of the sample can be measured in absolute or percentage terms, depending on whether one is trying to estimate a sample proportion (e.g., proportion of employees who rideshare) or an absolute number (e.g., average vehicle ridership).

For example, if the AVR value based on a survey of all employees is 1.35, and the AVR value based on a sample of employees is 1.45, the sample value would exceed or overestimate the population value, or true AVR, by 0.10 AVR points (i.e., $1.45 - 1.35 = 0.10$). Thus, the sampling error in this case would be 0.10 AVR points.

The size of the sampling error is critical when estimating the population AVR based on a sample. One of the most important and fundamental properties of random sampling is that it allows one to specify the sampling error before the actual sample is selected. Thus, the sample size can be selected according to the sampling error one is willing to accept.

7. Reliability

When choosing a sample, there is always the possibility of drawing a bad sample. The reliability or confidence of the sample is the risk that the sampling error may be larger than the one specified. This level of risk can be quantified and expressed in percentage terms when using a random sample. For example, when choosing a 95% reliability for a specified sampling error of 0.05 AVR, one would be 95% sure (or confident) that the sampling error will not exceed 0.05 AVR. When choosing a 99% reliability, one would be 99% sure (i.e., more confident) that the sampling error specified will not be exceeded. The higher the reliability of the sample, the higher the degree of accuracy of the estimated sampling values.

8. Sampling Techniques

a. Simple Random Sampling

While simple random sampling is not widely used in survey sampling, it constitutes the foundation of any other type of sampling procedure. Simple random sampling is a method of selecting a number of sampling units out of a population, such that each unit sampled has an equal chance of being selected. Thus, the sample is selected from a sampling frame on a purely random basis by using a random number table, such as the *Rand Corporation Random Number Table*, or by relying on a computer-generated

random selection process. Most statistical software designed for personal computers can do random selection from a data file.

Simple random sampling has one major shortcoming. By pure chance, the selection process might result in choosing too many or too few similar units with certain characteristics. The inclusion of a disproportionate number of employees from certain occupational, sex, or age categories would result in a poor sample or a sample that is not representative of the population of employees. For instance, a sample may contain too many professional employees who usually have low ridesharing participation rates. The inclusion of too many of these employees would generate a bad sample and a distorted picture of the employer's true AVR.

In addition, simple random sampling could be tedious and time consuming if done manually. Choosing random numbers from a random number table to pick a sample of 1,500 employees is not an attractive prospect. That is why sampling statisticians usually rely on simpler selection methods, such as systematic sampling.

b. Stratified Random Sampling

One way to compensate for the major shortcoming of simple random sample (i.e., the disproportionate selection of similar units) would be to arrange the population into relatively similar strata, categories, or layers. After the population is divided into non-overlapping groups, independent random samples may be selected proportionately from each stratum, or a systematic sample may be selected from the entire list.

For example, employers can stratify their employees by occupation, education, or age. After the population has been segregated into strata, a proportional sample can be selected from each stratum.

A proportional sample requires that the sample be allocated in proportion to the size of each stratum. Thus, larger strata would require larger samples, and smaller strata would require smaller samples. For example, if 40% of company employees reporting to the established window are professional employees, about 40% of the sample would have to be drawn from that stratum.

Thus, stratified random sample with proportional allocation is used to ensure the representativeness of the sample where it might otherwise be unlikely. Since one is interested in estimating AVR accurately, and travel modes may vary with occupation, a simple random sample may under represent certain employee occupations which may be in the majority. Stratification with proportional allocation would avoid this problem by grouping employees by occupation and ensuring that a proportionate sample is taken from every occupational group, large and small.

The advantage of stratification may not materialize unless the rationale for stratifying the population is logically related to the main objective of the survey. One would probably achieve little or no gain by stratifying the population of employees by height or weight,

when one is attempting to estimate AVR. On the other hand, it would be more logical to stratify employees by occupation because of the possible relationship between occupation and travel mode.

Thus, stratified random sampling is a refinement of probability sampling. The method of stratification still assures each item in the population a chance of being chosen in the sample. Stratification with proportional allocation would theoretically result in a more representative sample and in a more accurate AVR estimate. The sampling error is smaller when using stratified random sampling than when using simple random sampling.

c. Systematic Sampling

A systematic sampling approach can be used if the population is listed in an orderly way, such as alphabetical order, ascending or descending age, etc., or if the population is listed in a random order. This procedure uses a random starting point and a constant skip interval to select the sample from the population list.

For example, if the sampling frame contains 5,000 employee names and one needs a sample of 500, one would choose every 10th employee for the sample. To prevent any possible selection bias, one must choose the first employee at random. Thus, in this case, one would first pick a random number between 1 and 10, which would become the starting point in the selection process. Then, one would choose that first element and every 10th element thereafter. This method is known as systematic sampling with a random start.

In general, the results obtained from systematic sampling are approximately equal to the results obtained from simple random sampling. Thus, sampling error is about the same as for simple random sampling. In addition, since only the first element is chosen at random, it is simpler than simple random sampling, where all elements are selected at random. Systematic sampling is usually preferred to simple random sampling because of its expediency and simplicity in the selection process (Slonim, 1960).

E. Required Sample Size

While it is cheaper to obtain information from a random sample than from the entire population, it is more complicated since it requires knowledge of sampling theory and of the mechanics of sampling selection. As previously discussed, sampling also introduces sampling error in the estimation of relevant statistics, such as AVR values.

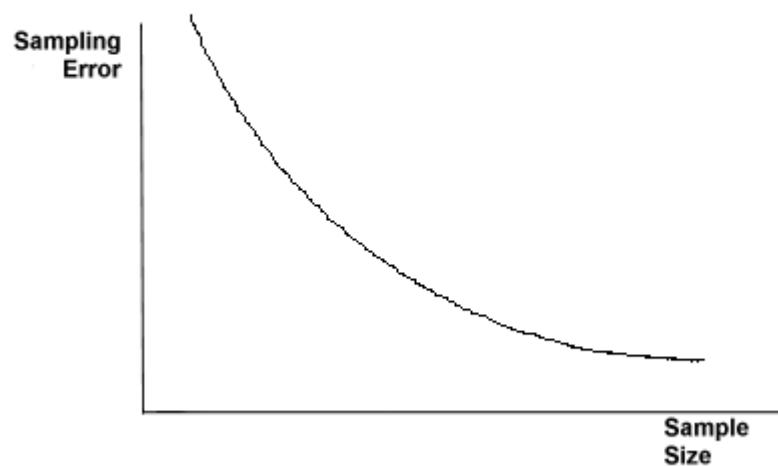
Thus, when selecting a sample, there is a trade off between sample size and the sampling error one is willing to accept (Figure 1). In general the larger the sample, the smaller the sampling error. For example, a large sample of employees would lead to an AVR value with small margin of error. By the same token, a small sample of employees would lead to an AVR value with a relatively larger margin of error.

1. Sampling Error and Reliability Criteria

Table 2 shows the sample sizes required to estimate AVR with a margin of error of plus or minus 0.025 AVR points. The reliability selected for this sampling error is 95%. This means that if the sample is properly selected, there will be a 95% chance that the sampling error will not exceed 0.025 AVR points. The sampling error and reliability have been selected after considerable discussion among Transportation Programs staff and consultation with an advisory panel composed of large employers subject to Rule 2202 (e.g., businesses, cities, and school districts), transportation management associations, and transportation demand management consultants.

Figure 1

Relation Between Sampling Error and Sample Size



Because AVR is simply the statistic used to evaluate employer's performance and progress toward the target AVR, the relatively small sampling error (i.e., 0.025 AVR) and high reliability (i.e., 95%) chosen are required to ensure the accuracy of the employer's reported AVR. This performance measure is also essential to gauge the SCAB's progress toward achieving the transportation performance standards of the federal and state Clean Air Acts.

2. Calculation of Sample Size

Employers choosing the sample method to estimate AVR must determine the appropriate sample size from Table 2. The ETC, or the person responsible for choosing the sample, must identify the category corresponding to the number of employees in the designated window. Then, the ETC must pick the required sample size from the second column. For example, if the number of window employees is 950, the corresponding sample size is 415.

Table 2
Minimum Sample Size Required to Estimate AVR *

Number of Employees in the Window	Size Sample
400 - 449	270
450 - 499	290
500 - 549	308
550 - 599	325
600 - 649	340
650 - 699	355
700 - 749	368
750 - 799	380
800 - 849	395
850 - 899	405
900 - 999	420
1,000 - 1,249	450
1,250 - 1,499	485
1,500 - 1,749	515
1,750 - 1,999	535
2,000 - 2,499	560
2,500 - 2,999	590
3,000 - 3,499	610
3,500 - 3,999	625
4,000 - 4,999	645
5,000 - 6,999	665
7,000 - 8,999	685
9,000 - 12,999	700
13,000 and Over	725

* Margin of error=0.025 AVR points; reliability=95%

3. Non-responses

Employers should monitor and minimize the non-responses to avoid bias in the sampling results. Several methods to control non-responses are explained in Kish (1965).

With the exception of a documented leave of absence, any sampled employee who does not respond to the AVR survey must be considered as an employee who drives alone to work. A minimum of 60% survey response rate is required.

F. The AQMD's Recommended Sampling Method

A sampling method that avoids the shortcomings of simple random sampling, and combines the advantages of stratified and systematic sampling, is systematic sampling with a random start, where the population is stratified according to a population characteristic which is relevant to the survey objective.

Given the relationship between occupation and employee modal choice, by stratifying the list of employees by occupation, a systematic sample would be spread over the entire population and will yield a fairly representative sample of employee mode split, as well as an accurate estimate of the employer's AVR.

As discussed earlier, when the list of sampling units is stratified according to population characteristics relevant to the survey objective, a systematic sampling yields more representative results than simple random sampling.

1. Drawing the Sample

The following steps outline the entire process of sample selection:

- a. Identify the population of window employees and compile the sampling frame. In other words, make a list containing the names of all employees arriving to work during the designated window.
- b. Assign an occupational category to each employee name (e.g., administrator, professional, technical, clerical, etc.).
- c. Sort the list of names by occupational categories to obtain blocks or strata of employees with similar occupation. As a result, the sampling frame will have all professional employees grouped together, all skilled employees grouped together, and so forth.
The position of these occupational groups in the list is irrelevant for sampling purposes. It does not matter whether professional employees are placed first, last, or somewhere in the middle of the list.
- d. Within each category or occupation, employee names must be sorted in alphabetical order. This would eliminate any patterns that may be related to travel mode and that may bias the sample.
- e. Determine the required sample size from Table 2
For example, if the number of window employees is 3,000, the sampling frame should contain 3,000 names arranged by occupation. Thus, the required sample size would be 600.
- f. Determine the skip interval to select a systematic sample of 600 names from the list of employees which has been stratified by occupation and arranged in alphabetical order within each occupational group, as described in (c) and (d).

The skip interval is calculated by dividing the population (3,000) by the required sample size (600):

$$\text{Skip interval} = 3,000 / 600 = 5.00$$

- g. Using a random number table, or any other suitable procedure, select a random number between 1 and 5. Assuming that the random number selected is 3 (it could have been 2 or 5 as well), the third name from the top of the list will be the first sampling unit selected or the first employee to be included in the sample.
- h. The second and subsequent names to be included in the sample are selected by adding the skip interval to the random start. For example:
Random start: 3
Pick the 3rd name from the top
 $3 + 5 = 8$: Pick the eighth name from the top.

The third and subsequent units will be selected by adding the skip interval to the last unit selected.

$8 + 5 = 13$	Pick the 13th name from the top;
$13 + 5 = 18$	Pick the 18th name from the top;
$18 + 5 = 23$	Pick the 23rd name from the top;
$23 + 5 = 28$	Pick the 28th name from the top;

The procedure must continue until the end of the list is reached. You will end up with a sample of 600 names.

2. Working with Inexact Skip Intervals

What happens if the skip interval is an inexact or fractional number? For example, if the number of employees in the window is 4,300, the required sample size is 633 ([Table 2](#)).

The skip interval is calculated by dividing the population (4,300) by the required sample size (633):

$$\text{Skip interval} = 4,300 / 633 = 6.79$$

Using a random number table, select a random number from 1 to 6.79. Since random number tables do not have decimals, you would pick a number between 1 and 679 and place the decimal point right after the first digit. Thus, if the random number you selected is 379, then the number to be used as the random start is 3.79. You will pick the third name from the top of the list.

The second and subsequent names to be included in the sample are selected by adding the skip interval (6.79) to the random start 3.79, as described in (h) above. The only difference is that when using fractional numbers to select the names from the list, you will round down to the smaller number or drop the decimal part. For example:

Random start: 3.79 Pick the 3rd name from the top (you have to drop the decimal)

Add the skip interval (6.79) to the random start (3.79) to get the second employee for the sample. You will continue to add the skip interval to pick the rest of your sample as follows:

3.79 + 6.79=10.58 Pick the 10th name from the top;
10.58 + 6.79=17.37 Pick the 17th name from the top;
17.37 + 6.79=24.16 Pick the 24th name from the top;

The procedure must be continued until the end of the employee list is reached. When you get to the end of the list, you will have selected 633 employees for your sample.

To learn about other procedures to deal with inexact skip intervals, see Kish (1960), pp. 116-117.

3. Multi-Site Employers

Employers with multi-site or geographic plans have the option of aggregating AVR for worksites located within the same AVR target area. These employers also have the option to use a random sample of employees in any site with at least 400 employees in the window. If the employer chooses the random sample option, independent random samples must be drawn from each site according to the procedures outlined in this Section, or in Section G if choosing other than the AQMD's recommended approach.

When aggregating worksite AVRs calculated with a random sample, the sample AVRs must be weighted by the population of employees in the window. For example, if an employer decides to aggregate the sample AVRs for two different worksites, the aggregate AVR must be calculated as follows:

$$AVR = \frac{AVR1 * N1 + AVR2 * N2}{N1 + N2}$$

Where AVR1 and AVR2 are the sample AVRs for sites 1 and 2, and N1 and N2 are the number of employees in the window for sites 1 and 2, respectively.

Thus, if there are two worksites and their sample AVRs are 1.2 and 1.6 respectively; and the number of employees in the windows are 1,000 and 3,000 respectively, the aggregate AVR must be computed as follows:

$$AVR = \frac{1.2 * 1,000 + 1.6 * 3,000}{1,000 + 3,000} = 1.5$$

G. Exceptions to the AQMD's Recommended Sampling Method

Employers have the option of proposing a sampling method other than the AQMD's recommended approach. The proposed sampling method must be a stratified random sampling as described on p. 9 of this report. A simple random sample or systematic approach may, in turn, be used to select employees proportionately from each stratum.

In selecting a stratification variable, the employer should be concerned primarily with those that are presumably associated with the variables that the company wishes to represent accurately. In social science research, socio-demographic variables (e.g., age, gender, education, and occupation) are often used because they are related to many other variables (Babbie, 1973) and are likely to affect individual behavior.

Simple random sampling or systematic sampling approaches will not be accepted unless they have a stratification procedure built into them.

The sample size must be obtained from Table 2.

H. Requesting Approval to Use a Random Sample

Employers who plan to use the sampling method to estimate AVR must seek approval and certification from the AQMD at least 3 months before the AVR survey is conducted.

All requests must be submitted in writing to the AQMD and must be addressed as follows:

AQMD
Transportation Programs
Request for Approval of Sampling Method
21865 Copley Drive
Diamond Bar, CA 91765

Sampling methodologies will be evaluated and approved, or disapproved, within 30 days from the date the request is received by the AQMD.

Employers must not deviate from sampling methodologies already approved by the AQMD. Modifications to sampling methodologies already approved must be requested in writing to the AQMD and will be subject to the complete certification process again.

1. AQMD's Method

Employers who select the AQMD's recommended sampling approach must include the following information in their written request to the AQMD:

- Commitment to use the AQMD's method;
- Number of employees in the window; and
- The sample size to be used in the survey.

Employers who make a commitment to use the AQMD's sampling method but fail to carry it out, as described in Section F of the guidelines, will be required to re-take the sample and to adhere to the AQMD's prescribed procedures.

2. Other than the AQMD's Method

Employers using a sampling methodology different from the AQMD's recommended approach must include the following information in their written request to the AQMD:

- Justification for alternative methodology;
- A commitment to use the proposed method if approved by the AQMD;
- Number of employees in the window;
- The sample size to be used in the survey; and
- A detailed description of the sampling method, including the stratification procedure to be used.

The detailed description of the sampling methodology must include a step-by-step approach similar to the process illustrated in Section F of the guidelines.

I. Required Sample Documentation

To ensure that the sampling procedures used by employers are valid, certain information must be submitted by employers to document how the sampling methodology has been executed. At a minimum, the following items must be submitted with the employee commute reduction program:

- Complete documentation on the sampling frame, i.e., the number of employees stratified by the variable of choice;
- The sample size with a complete description of the mechanics of sampling selection, i.e., random starting point, skip interval and number of sampling units selected from each stratum or category;
- The number of employees not responding to the survey;
The number of non-respondents (a non-response must be counted as a solo driver); a minimum of 60% response rate is required;

- While the names of employees in the sampling frame and the names of sampled employees do not have to be submitted to the AQMD, they must be properly stored and made available upon request;
- A table comparing sample with population data;
This can be accomplished by comparing the sample and population proportions of the various categories used to stratify the population. This information is crucial and will be used by Transportation Programs staff to determine whether the sample is fairly representative of the population; and
- Employers using the AQMD's method must compare the occupational characteristics of the sample with the occupational characteristics of the population of employees. Table 3 shows how the comparison should be made and the type of information that must be submitted to the AQMD for validation purposes.

Large differences between sample and population proportions will be viewed as indicative that the sample is not representative of the population. These large differences will be considered sufficient evidence to render the sample invalid and to reject it on statistical grounds.

When employee occupation is the stratification variable, the AQMD will consider a sample to be representative of the employee population if the maximum difference between the sample and population proportions does not exceed 5% for any of the occupational categories used in the sample.

Table 3 presents a hypothetical case where the sample is considered to be representative of the population. As shown by the shaded columns of the table, the differences between the population and sample proportions are relatively small. The largest difference is found in the Professionals category: 20% of the employees in the population are "Professional" as compared with 15% of the employees in the sample. Since in this case the difference does not exceed 5%, the sample is considered to be representative of the population.

In contrast, Table 4 presents a hypothetical case where 12% of the employees in the population are "Official/Administrators" as compared with 18% of the employees in the sample. In this case, official and administrative personnel are overrepresented in the sample. Since the difference between the sample and population proportion for this particular occupation group is greater than 5%, the sample would not be considered to be representative of the population, and would be rejected.

Employers using other than the AQMD's recommended sampling approach must compare the sample with the population data following a format similar to Tables 3 or 4. In this case, however, employers will use the stratification variable of their choice.

Table 3
Comparison Between Population and Sample Proportions:
Example of **ACCEPTABLE** Differences (Hypothetical Case)

Employee Occupation	Population Data		Sample Data	
	Number	Percent	Number	Percent
Official/Administrators	420	12.0%	105	15.0%
Professional	700	20.0%	133	19.0%
Technical	1400	40.0%	245	35.0%
Office/Clerical	245	7.0%	63	9.0%
Skilled crafts	350	10.0%	84	12.0%
Service /Maintenance	0	0.0%	0	0.0%
Other 1	175	5.0%	21	3.0%
Other 2	210	6.0%	49	7.0%
TOTAL	3,500	100.0%	700	100.0%

Population data include window employees stratified by occupation.
Sample data include sampled employees classified by occupation.

Table 4
Comparison Between Population and Sample Proportions:
Example of **UNACCEPTABLE** Differences (Hypothetical Case)

Employee Occupation	Population Data		Sample Data	
	Number	Percent	Number	Percent
Official/Administrators	420	12.0%	126	18.0%
Professional	700	20.0%	133	19.0%
Technical	1400	40.0%	245	35.0%
Office/Clerical	245	7.0%	63	9.0%
Skilled crafts	350	10.0%	84	12.0%
Service /Maintenance	0	0.0%	0	0.0%
Other 1	175	5.0%	14	2.0%
Other 2	210	6.0%	35	5.0%
TOTAL	3,500	100.0%	700	100.0%

Population data include window employees stratified by occupation.
Sample data include sampled employees classified by occupation.

Appendix I - Statistical Formulas Used to Determine Sample Sizes

Sampling Without Replacement:

$$E^2 = Z^2 * S^2/n * (N-n)/N$$

E: Sampling error;

Z: Sample reliability (Z=1.96 for 95% reliability);

S²: Estimated population variance;

N: Population size; and

n: Sample size

Thus, sample size (n) is obtained as follows:

$$n = (N * Z^2 * S^2) / (N * E^2 + Z^2 * S^2)$$

For example. If:

E= 0.025 AVR points;

Z= 1.96;

S²= 0.1197; and

N= 5,000

n= 642

See Yamane (1967), Ch.5.

Appendix II - References

Earl Babbie. 1986. The Practice of Social Research. Belmont, California: Wadsworth Publishing Co.

Earl Babbie. 1973. Survey Research Methods. Belmont, California: Wadsworth Publishing Co.

Hubert Blalock, Jr. 1979. Social Statistics. New York: McGraw-Hill Book Co.

Fred Kerlinger. 1978. Foundations of Behavioral Research. New York: Holt, Rinehart, and Winston.

Leslie Kish. 1967. Survey Sampling. New York: John Wiley and Sons, Inc.

Richard L. Scheaffer, W. Mendenhall, and L. Ott. 1986. Elementary Survey Sampling. Boston: Duxbury Press.

Morris J. Slonim. 1960. Sampling in a Nutshell. New York: Simon and Schuster.

Stevens, W. F. 1990. Improving the Effectiveness of Ridesharing Programs. Transportation Quarterly 44, 4: 563-78.

Taro Yamane. 1967. Elementary Sampling Theory. Englewood Cliffs, N.J.: Prentice Hall, Inc.